Economics Department

Market Microstructure Models and the Markov Property

João Amaro de Matos
and
Marcelo Fernandes

ECO No. 2000/19

EUI WORKING PAPERS

EUROPEAN UNIVERSITY INSTITUTE

EUROPEAN UNIVERSITY INSTITUTE, FLORENCE

ECONOMICS DEPARTMENT

EUI Working Paper **ECO** No. 2000/19

Market Microstructure Models and the Markov Property

JOÃO AMARO DE MATOS
and
MARCELO FERNANDES

BADIA FIESOLANA, SAN DOMENICO (FI)

Printed in Italy in September 2000
European University Institute
Badia Fiesolana
I – 50016 San Domenico (FI)
Italy

Market Microstructure Models and the Markov Property

João Amaro de Matos
Faculdade de Economia
Universidade Nova de Lisboa

and

Marcelo Fernandes
Department of Economics
European University Institute

July 2000

Abstract

This paper develops a framework to test alternative market microstructure models of the bid-ask spread. If, on the one hand, information-based models result in bid and ask quotes that are non-Markovian, on the other hand, the Markov property may hold in equilibrium settings where the market maker serves as an intermediary. We thus derive a simple nonparametric test for Markovian dynamics, suitable to high frequency data, so as to address the merits of information-based and equilibrium models. Finally, we examine whether or not bid-ask spreads follow Markov processes using data from the New York Stock Exchange.

1 Introduction

Information-based models in the market microstructure literature use adverse selection arguments to show how, even in competitive markets without explicit transaction costs, bid-ask spreads would exist. The main intuition behind these models dates back to Baghehot (1971), who claims that the market maker sets a bid-ask spread to balance losses on trades with insider (or informed) agents with gains on trades with liquidity (or uninformed) traders. This framework allows for the examination of market dynamics and hence provides insights into the adjustment process of prices. Easley and O'Hara (1992) are the first to delineate the link between the existence of information, the timing of trades and the stochastic processes of the bid and ask prices. In particular, because time is not exogenous to the formation of prices, bid and ask quotes cannot follow Markov processes.

Another way to explain the existence of bid-ask spreads rests on equilibrium models. Amaro de Matos and Rosário (2000) extend the framework of Platen and Rebolledo (1996) to introduce intermediation in financial markets. Taking advantage of their market power, the market makers set bid-ask spreads in order to maximize their profits. According to the nature of the demand and supply processes for the underlying security, bid and ask quotes may follow Markov processes.

This paper develops a nonparametric test for the Markov property that is particularly suitable to high frequency data, essential to empirical microstructure studies (Goodhart and O'Hara, 1997). This allows for testing the Markovian nature of the bid and ask prices so as to address the relative merits of information-based and equilibrium models. Any evidence supporting a Markovian character of bid and ask prices invalidates information-based models, though it is consistent with equilibrium models. In contrast, rejecting the Markov property does not rule out equilibrium models as interpretative tools. Indeed, these models could also generate non-Markovian processes for the bid and ask prices.

An empirical application is performed using data from five stocks actively traded on the New York Stock Exchange (NYSE), namely Boeing, Coca-Cola, Disney, Exxon, and IBM. This market has been chosen since it is well known for strong adverse selection costs as opposed, for instance, to the London Stock Exchange (Snell and Tonks, 1998). In that sense, not rejecting the null hypothesis of the Markov property with these data provides credible evidence against information-based models. The results indicate that the Markov assumption is consistent with the Coca-Cola, Disney and Exxon bid-ask spreads, whereas the converse is true for Boeing and IBM.

The remainder of this paper is organized as follows. Section 2 describes the differences between the asymmetric information and general equilibrium approaches to modeling the bid-ask spread. It turns out that the main testable distinction refers to whether or not the bid-ask spread follows a Markov process. Section 3 starts by discussing how to design a nonparametric test for Markovian dynamics which is suitable to high frequency data. Next, we show the asymptotic normality of the test statistic both under the null hypothesis that the Markov property holds and under a sequence of local alternatives. Section 4 applies the above ideas to the NYSE high-frequency data to test for adverse selection. Section 5 summarizes the results and offers some concluding remarks. For ease of exposition, we collect all proofs and technical lemmas in the appendix.

2 Market microstructure models

2.1 Information-based models

To concentrate on the effect of information on prices, Easley and O'Hara (1992) assume a single market maker who is risk neutral and acts competitively. The first characteristic rules out direct inventory effects, whilst the latter implies the existence of at least potential competitors. Let V denote the value of a certain asset and define an information event as the occurrence of a signal ψ about V. The signal can take on one of two values, L and H, with probabilities $\delta > 0$ and $1 - \delta > 0$. The expected value of the asset conditional on the signal is $E(V \mid \psi = L) = V_L$ or $E(V \mid \psi = H) = V_H$. If no information event occurs ($\psi = 0$), then the expected value of the asset simply remains at its unconditional level $V_* = \delta V_L + (1 - \delta) V_H$.

Information events occur with probability $\alpha \in (0,1)$ before the start of the current trading day. This uncertainty reflects the fact that, since uninformed market participants do not receive any signals, they may also not know whether any new information even exists (e.g. Dow-Jones Rumor Wire). Trade can arise from uninformed and/or informed traders. To keep the focus on information effects, Easley and O'Hara assume that the informed traders are risk neutral and price takers so as to rule out strategic behavior. As such, the resulting trading strategy reads: If a high signal occurs, an insider will buy the stock if the current quote is below V_H; if a low signal occurs, then he will sell if the quote is above V_L.

To avoid no-trade equilibria, some uninformed market participants must trade for non-speculative reasons such as liquidity needs or portfolio considerations. Further, suppose that there is a fraction γ of potential sellers and a fraction $1 - \gamma$ of potential buyers among liquidity traders. Uninformed buyers trade with probability ϵ_B, whereas an uninformed seller's trading probability is ϵ_S.

The assumptions of risk neutrality and competitive behavior for the market maker dictate that price quotes yield zero expected profit conditional on a trade at the quote. As insider traders will profit at the market maker's expense, the probability $\mu \in (0,1)$ that a trade is actually information-based is crucial for determining the price quotes. There are several ways to interpret μ: e.g. fraction of the trader population observing the signal, probability of information disclosure to the selected trader, and exogenous order arrival rate.

Insiders always trade provided that prices are not at their full information value, whereas liquidity traders may buy, sell or not trade according to their type (buyer or seller) and motivation. Therefore, non-trading occurs only when an uninformed trader checks the quotes and decides for portfolio reasons (as captured by ϵ_B and ϵ_S) not to trade. This can happen both when there has been an information event and when there has not. Of course, no-trade outcomes are more likely to occur when there is no new information than when new information arrives:

$$(1-\gamma)(1-\epsilon_B)+\gamma(1-\epsilon_S) > (1-\mu)[(1-\gamma)(1-\epsilon_B)+\gamma(1-\epsilon_S)]. \quad (1)$$

The tree diagram in Figure 1 summarizes the trading process.

The market maker knows the structure of the market and updates his beliefs by Bayes rule. It is precisely this revision that causes quotes, and thus prices, to adjust. The quote-setting process for the first trade of the day unfolds as follows. To determine the evolution of beliefs, notice that there is no need to keep track of the order in which sales, buys and no-trades arrive in the market. Suppose, for instance, that in the past t trading intervals there were n_t no-trades, β_t buys, and s_t sales. Then, (n_t, β_t, s_t) is a sufficient statistic for Q^t. For instance, the probability the market maker assigns to the absence of new information given this trading history reads

$$
\begin{aligned}
\Pr(\psi = 0 \mid Q^t) \;&=\; (1-\alpha)(\gamma\epsilon_S)^{s_t}[(1-\gamma)\epsilon_B]^{\beta_t} \Big\{(1-\alpha)(\gamma\epsilon_S)^{s_t}[(1-\gamma)\epsilon_B]^{\beta_t} \\
&\quad + (1-\mu)^{n_t}\Big[\alpha\delta(\mu + (1-\mu)\gamma\epsilon_S)^{s_t}((1-\mu)(1-\gamma)\epsilon_B)^{\beta_t} \\
&\quad + \;\alpha(1-\delta)((1-\mu)\gamma\epsilon_S)^{s_t}(\mu + (1-\mu)(1-\gamma)\epsilon_B)^{\beta_t}\Big]\Big\}^{-1}.
\end{aligned}
$$

More importantly, as beliefs depend on (n_t, β_t, s_t), quotes will also depend on these quantities. In particular, the bid and ask prices at time $t+1$ can be written as

$$
\begin{aligned}
b_{t+1} \;&=\; \Pr(\psi = L \mid n_t, \beta_t, s_t + 1)V_L + \Pr(\psi = 0 \mid n_t, \beta_t, s_t + 1)V_* \\
&\quad + \Pr(\psi = H \mid n_t, \beta_t, s_t + 1)V_H \qquad (2) \\
a_{t+1} \;&=\; \Pr(\psi = L \mid n_t, \beta_t + 1, s_t)V_L + \Pr(\psi = 0 \mid n_t, \beta_t + 1, s_t)V_* \\
&\quad + \Pr(\psi = H \mid n_t, \beta_t + 1, s_t)V_H, \qquad (3)
\end{aligned}
$$

respectively.

As is apparent, both bid and ask quotes depend not only on the most recent quote, but also on the total numbers of previous buys, sales and no-trade outcomes. This means that they do not satisfy the Markov property as opposed to the process of trade outcomes (n_t, β_t, s_t).

2.2 Equilibrium Models

A different approach for explaining the behavior of the bid-ask spread rests on financial market equilibrium conditions in the presence of financial intermediation. The main duty of market makers is to provide liquidity (O'Hara, 1995). Taking advantage of their access to superior information on the trade orders, the market makers set different prices for buys and sells so as to profit from the spread.

Amaro de Matos and Rosário (2000) begin by specifying the stochastic processes for the demand and supply of a financial asset as in Platen and Rebolledo (1996). In the absence of intermediation, the demand process reads

$$D\left(t, p_t, \rho_t^d\right) = \alpha\, t - \beta^d p_t + \gamma^d \rho_t^d, \tag{4}$$

where $\left(\alpha, \beta^d, \gamma^d\right) \geq 0$, p_t is the security's price and ρ_t^d denotes the cumulative demand. The linear trend accounts for a deterministic amount of the security that is traded.

The stochasticity of the demand stems from the fact that the dynamics of ρ_t^d obey the following stochastic differential equation

$$\mathrm{d}\rho_t^d = \lambda^d(\overline{p}_t - p_t)\,\mathrm{d}t + \sigma_t^d(p_t)\,\mathrm{d}W_t, \tag{5}$$

where λ^d is non-negative, so as to reflect investors' desire to buy low and sell high. The price $\overline{p}_t$ corresponds to the risk-neutral valuation of the security. The model's source of noise is a standard Brownian motion W_t associated with a diffusion function $\sigma_t^d(p_t)$ that may depend on the price p_t and time t. Notice moreover that it is this particular specification of ρ_t^d that determines the Markovian nature of the demand process.

The supply process is modeled in a similar way to the demand, viz.

$$S\left(t, p_t, \rho_t^s\right) = \alpha\, t + \beta^s p_t + \gamma^s \rho_t^s, \tag{6}$$

where $\left(\beta^d, \gamma^d\right) \geq 0$ and the cumulative supply ρ_t^s follows

$$\mathrm{d}\rho_t^s = \lambda^s(p_t - \overline{p}_t)\,\mathrm{d}t + \sigma_t^s(p_t)\,\mathrm{d}W_t. \tag{7}$$

In the absence of intermediation, the market-clearing condition imposing the equality between the demand and supply processes determines a single price governed by a Markov process as in Platen and Rebolledo (1996).

The presence of M market makers implies that the demand and supply do not interact directly. In fact, intermediaries use their market power to set a traded amount q_t that maximizes their profit. They will therefore sell at a price a_t such that $q_t = D\left(t, a_t, \rho_t^d\right)$ and buy at a price b_t satisfying $q_t = S\left(t, b_t, \rho_t^s\right)$. A market maker then maximizes the profit function

$$V(q_t) = \max\left\{q_t\left(a_t - b_t\right) - K(q_t)\right], 0\}. \tag{8}$$

Assuming a positive, increasing cost function of the form $K(q_t) = c_0 + c_1 q_t + c_2 q_t^2$, Amaro de Matos and Rosário (2000) show that, in equilibrium, the bid and ask prices are, respectively

$$\begin{aligned}
a_t &= \frac{\alpha\left(A - \beta^d - \beta^s\right)}{\beta^d\left[A + (M-1)\left(\beta^d + \beta^s\right)\right]} t + \frac{\left[A + (M-1)\beta^d - \beta^s\right]\gamma^d}{\beta^d\left[A + (M-1)\left(\beta^d + \beta^s\right)\right]} \rho_t^d \\
&\quad - \frac{M\beta^d\gamma^s}{\beta^d\left[A + (M-1)\left(\beta^d + \beta^s\right)\right]} \rho_t^s + \frac{Mc_1\beta^d\beta^s}{\beta^d\left[A + (M-1)\left(\beta^d + \beta^s\right)\right]}, \\
b_t &= -\frac{\alpha\left(A - \beta^d - \beta^s\right)}{\beta^s\left[A + (M-1)\left(\beta^d + \beta^s\right)\right]} t + \frac{M\beta^s\gamma^d}{\beta^s\left[A + (M-1)\left(\beta^d + \beta^s\right)\right]} \rho_t^d \\
&\quad - \frac{\left[A + (M-1)\beta^s - \beta^d\right]\gamma^s}{\beta^s\left[A + (M-1)\left(\beta^d + \beta^s\right)\right]} \rho_t^s - \frac{Mc_1\beta^d\beta^s}{\beta^s\left[A + (M-1)\left(\beta^d + \beta^s\right)\right]},
\end{aligned}$$

where $A = 2(\beta^d + \beta^s + c_2\beta^d\beta^s)$. In particular, this implies that the bid-ask spread will follow a diffusion process. It is interesting to note that, in the limit case where the number of market makers is arbitrarily large ($M \to \infty$), the bid-ask spread is constant and equal to c_1.

3 Testing for Markovian dynamics

Despite the innumerable studies rooted in Markov processes, there are only two testing procedures available in the literature: Aït-Sahalia (1997) and Fernandes and Flôres (1999). To build a nonparametric testing procedure, the first uses the fact that the Chapman-Kolmogorov equation must hold in order for a Markov process compatible with the data to exist. If, on the one hand, the Chapman-Kolmogorov representation involves a quite complicated nonlinear functional relationship among transition probabilities of the process, on the other hand, it brings about several advantages. First, estimating transition distributions is straightforward and does not require any prior parameterization of conditional moments. Second, a test based on the whole transition density is obviously preferable to tests based on specific conditional moments. Third, the Chapman-Kolmogorov representation is well defined, even within a multivariate context.

Fernandes and Flôres (1999) develop alternative ways of testing whether discretely recorded observations are consistent with an underlying Markov process. Instead of using the highly nonlinear functional characterization provided by the Chapman-Kolmogorov equation, they rely on a simple characterization out of a set of necessary conditions for Markov models. As in Aït-Sahalia (1997), the testing strategy boils down to measuring the closeness of density functionals which are nonparametrically estimated by kernel-based methods.

In principle, to verify which microstructure model better approximates the market reality, one may apply the above testing procedures to check whether the bid-ask spread follows a Markov process. However, these tests are built under the assumption that the data are equally spaced in time, which does not hold for high frequency data. As such, market microstructure analyses call for a nonparametric test of the Markov assumption, which is suitable to data irregularly spaced in time. For that purpose, we build on the theory of subordinated Markov processes in which a continuous-time strong Markov process is observed only when it crosses some discrete level. Such a sampling scheme accommodates not only the irregular spacing of transaction data, but also price discreteness.

Let t_i $(i = 1, 2, \ldots)$ denote the observation times of the continuous-time price process $\{X_t,\ t > 0\}$ and assume that $t_0 = 0$. Suppose further that the shadow price follows a strong Markov process. To account for price discreteness, we assume that prices are observed only when the cumulative change in the shadow price is at least c, say a basic tick. The price duration then reads

$$d_{i+1} = t_{i+1} - t_i = \inf_{\tau>0} \{|X_{t_i+\tau} - X_{t_i}| \geq c\}. \tag{9}$$

The data available for statistical inference are the discrete realizations of the price process $(X_1, \ldots, X_n)$, where $X_i = X_{t_i}$, and their correspondent durations $(d_1, \ldots, d_n)$.

The observation times $\{t_i,\ i = 1, 2, \ldots\}$ form a sequence of increasing stopping times of the continuous-time Markov process $\{X_t,\ t > 0\}$, hence the discrete-time process $\{X_i,\ i = 1, 2, \ldots\}$ satisfies the Markov property as well. Further, the duration d_{i+1} between t_{i+1} and t_i is a measurable function of the path of $\{X_t,\ 0 < t_i \leq t \leq t_{i+1}\}$, and thus depends on the information available at time t_i only through X_i, i.e. $d_i \perp\!\!\!\perp d_j \mid X_i$ for every $0 < j < i$ in the notation of Dawid (1979). Therefore, we test the Markov assumption by checking the property of conditional independence between consecutive durations given the current price realization.

The null hypothesis of conditional independence implied by the Markov character of the price process then reads

$$H_0: \quad f_{iXj}(a_1, x, a_2) = f_{i|X}(a_1) f_{Xj}(x, a_2) \ \text{ a.s.,} \tag{10}$$

where f_{iXj}, $f_{i|X}$ and f_{Xj} denote the joint density of (d_i, X_i, d_j), the conditional density of d_i given X_i and the joint density of (X_i, d_j), respectively. To keep the nonparametric nature of the testing procedure, we employ kernel smoothing to estimate both the right- and left-hand sides of (10). Next, it suffices to gauge how well the density restriction in (10) fits the data by the means of some discrepancy measure.

For the sake of simplicity, we consider the mean squared difference, yielding the following test statistic

$$\Lambda_f = E[f_{iXj}(d_i, X_i, d_j) - f_{i|X}(d_i|X_i) f_{Xj}(X_i, d_j)]^2. \tag{11}$$

The sample analog is then

$$\Lambda_{\hat{f}} = \frac{1}{n-i+j} \sum_{k=1}^{n-i+j} \left[\hat{f}_{iXj}(d_{k+i-j}, X_{k+i-j}, d_k) - \hat{g}_{iXj}(d_{k+i-j}, X_{k+i-j}, d_k)\right]^2,$$

where $\hat{g}_{iXj}(d_{k+i-j}, X_{k+i-j}, d_k) = \hat{f}_{i|X}(d_{k+i-j}|X_{k+i-j})\hat{f}_{Xj}(X_{k+i-j}, d_k)$. Any other evaluation of the integral on the right-hand side of (11) can be used.

At first glance, deriving the limiting distribution of $\Lambda_{\hat{f}}$ seems to involve a number of complex steps since one must deal with the cross-correlation among $\hat{f}_{iXj}$, $\hat{f}_{i|X}$ and $\hat{f}_{Xj}$. Happily, the fact that the rates of convergence of the three estimators are different simplifies things substantially. In particular, $\hat{f}_{iXj}$ converges slower than $\hat{f}_{i|X}$ and $\hat{f}_{Xj}$ due to its higher dimensionality. As such, estimating the conditional density $f_{i|X}$ and the joint density f_{Xj} does not play a role in the asymptotic behavior of the test statistic. The following proposition uses the tools provided by Aït-Sahalia (1994) to demonstrate the asymptotic normality of the standardized test statistic.

PROPOSITION 1: *For $u, v \in \mathbb{R}^3$, define $e_K \equiv \int K^2(u)\,\mathrm{d}u$ and $v_K \equiv \int \left[\int K(u)K(u+v)\,\mathrm{d}u\right]^2 dv$. Under the null and suitable regularity conditions (Aït-Sahalia, 1994), the statistic*

$$\hat{\lambda}_n = \frac{n\, b_n^{1/2} \Lambda_{\hat{f}} - b_n^{-1/2} \hat{\delta}_\Lambda}{\hat{\sigma}_\Lambda} \xrightarrow{d} N(0,1),$$

where $b_n = b_{d,n}^2\, b_{x,n}$ is the bandwidth for the kernel estimation of the joint density f_{iXj}, and $\hat{\delta}_\Lambda$ and $\hat{\sigma}_\Lambda^2$ are consistent estimates of $\delta_\Lambda = e_K\, E(f_{iXj})$ and $\sigma_\Lambda^2 = v_K\, E(f_{iXj}^3)$, respectively.

Consider now the following sequence of local alternatives

$$H_1^{[n]}: \ \sup \left| f_{iX_j}^{[n]}(a_1, x, a_2) - g_{iX_j}^{[n]}(a_1, x, a_2) - \epsilon_n \ell(a_1, x, a_2) \right| = o(\epsilon_n), \tag{12}$$

where $\left\| f_{iX_j}^{[n]} - f_{iX_j} \right\| = o\left(n^{-1} b_n^{-1/2}\right)$, $\epsilon_n = n^{-1/2} b_n^{-1/4}$ and $\ell(\cdot, \cdot, \cdot)$ is such that $E[\ell(a_1, x, a_2)] = 0$ and $\ell_2 \equiv E[\ell^2(a_1, x, a_2)] < \infty$. The next result illustrates the fact that the testing procedure entails nontrivial power under local alternatives that shrink to the null at rate ϵ_n.

PROPOSITION 2: *Under the sequence of local alternatives $H_1^{[n]}$ and assumptions A1 to A4,* $\hat{\lambda}_n \xrightarrow{d} N\left(\ell_2/\sigma_\Lambda, 1\right)$.

4 Empirical exercise

We focus on New York Stock Exchange (NYSE) data ranging from September to November 1996. In particular, we look at five actively traded stocks from the Dow Jones index: Boeing, Coca-Cola, Disney, Exxon, and IBM.[1] Trading on the NYSE is organized as a combined market maker/order book system. A designated specialist composes the market for each stock by managing the trading and quoting processes and providing liquidity. Apart from an opening auction, trading is continuous from 9:30 to 16:00. Spread durations are defined as the time interval

[1] Data were kindly provided by Luc Bauwens and Pierre Giot and refer to the NYSE's Trade and Quote (TAQ) database. Giot (1999) describes the data more thoroughly.

needed to observe a change either in the bid or in the ask price.

For all stocks, durations between events recorded outside the regular opening hours of the NYSE, as well as overnight spells, are removed. As documented by Giot (1999), durations feature a strong time-of-day effect related to predetermined market characteristics, such as trade opening and closing times and lunch time for traders. To account for this anomaly, we also consider seasonally adjusted spread durations $d_i^* = d_i/\phi(t_i)$, where d_i is the original spread duration in seconds and $\phi(\cdot)$ denotes a time-of-day factor determined by averaging durations over thirty-minutes intervals for each day of the week and fitting a cubic spline with nodes at each half hour.

The motivation for working with the bid-ask spread rather than bid and ask prices is simple. The results reported in Table I show that the bid and ask quotes are both integrated of order one, and hence non-stationary. In contrast, there is no evidence of unit roots in the bid-ask spread process. As kernel density estimation relies on the assumption of stationarity, spread data are thus more convenient as input for the subsequent analysis.

All density estimations are carried out using a (product) Gaussian kernel, namely

$$K(u) = (2\pi)^{-3/2} \exp\left(-\frac{u_1^2 + u_2^2 + u_3^2}{2}\right), \tag{13}$$

which implies that $e_K = (4\pi)^{-3/2}$ and $v_K = (8\pi)^{-3/2}$. Bandwidths are chosen according to Silverman's (1986) rule of thumb adjusted so as to conform to the required degree of undersmoothing (see Aït-Sahalia, 1994).

Table II reports mixed results in the sense that the Markov hypothesis seems to suit only some of the bid-ask spreads under consideration. Clear rejection is detected in the Boeing, Coca-Cola and IBM bid-ask spreads, indicating that adverse selection may play a role in the formation of their prices. In contrast, there is no indication of non-Markovian behavior in the Disney and Exxon bid-ask spreads. These results agree to some extent with Fernandes and Grammig's (2000) analysis, which identifies significant asymmetric information effects only in the Boeing and IBM price durations.

5 Conclusion

This paper has developed a test to check whether information-based models of market microstructure fit well data on bid-ask prices as compared to equilibrium models. The testing strategy rests on the fact that information-based models result in bid-ask spreads that are non-Markovian, whereas the Markov property may hold only in equilibrium models. We derive a simple nonparametric procedure, which is particularly suitable to high frequency data, to test the Markov assumption.

Using data from the New York Stock Exchange, we show that whether the Markovian hypothesis is reasonable is indeed an empirical issue. A Markovian character suits the Disney and Exxon bid-ask spreads well, thus providing evidence in support of equilibrium models. In contrast, the rejection of the Markov assumption for Boeing, Coca-Cola and IBM does not permit one to discriminate between information-based and equilibrium models.

Appendix: Proofs

PROOF OF PROPOSITION 1: Consider the second-order functional Taylor expansion of the test statistic

$$\Lambda_{f+h} = \Lambda_f + \mathrm{D}\Lambda_f(h) + \frac{1}{2}\,\mathrm{D}^2\Lambda_f(h,h) + O\left(||h||^3\right),$$

where h denotes the perturbation $h_{iXj} = \hat{f}_{iXj} - f_{iXj}$. Under the null hypothesis that $f_{iXj} = g_{iXj}$, both Λ_f and $\mathrm{D}\Lambda_f$ equal zero. To appreciate the singularity of the latter, it suffices to compute the Gâteaux derivative of $\Lambda_{f,h}(\lambda) = \Lambda_{f+\lambda h}$ with respect to λ evaluated at $\lambda \to^{+} 0$. Let

$$g_{iXj}(\lambda) = \frac{\int [f_{iXj} + \lambda h_{iXj}](a_1, x, a_2)\mathrm{d}a_2 \int [f_{iXj} + \lambda h_{iXj}](a_1, x, a_2)\mathrm{d}a_1}{\int [f_{iXj} + \lambda h_{iXj}](a_1, x, a_2)\mathrm{d}(a_1, a_2)}.$$

It follows then that

$$\begin{aligned}\frac{\partial \Lambda_{f,h}(0)}{\partial \lambda} &= 2\int [f_{iXj} - g_{iXj}][h_{iXj} - \mathrm{D}g_{iXj}] f_{iXj}(a_1, x, a_2)\,\mathrm{d}(a_1, x, a_2)\\ &\quad + \int [f_{iXj} - g_{iXj}]^2 h_{iXj}(a_1, x, a_2)\,\mathrm{d}(a_1, x, a_2),\end{aligned}$$

where $\mathrm{D}g_{iXj}$ is the functional derivative of g_{iXj} with respect to f_{iXj}, namely

$$\mathrm{D}g_{iXj} = \left(\frac{h_{iX}}{f_{iX}} + \frac{h_{Xj}}{f_{Xj}} - \frac{h_X}{f_X}\right) g_{iXj}.$$

As is apparent, imposing the null hypothesis induces singularity in the first functional derivative $\mathrm{D}\Lambda_f$. It is easy to see that, under the null, the second-order derivative reads

$$\mathrm{D}^2\Lambda_f(h,h) = 2\int [h_{iXj}(a_1, x, a_2) - \mathrm{D}g_{iXj}(a_1, x, a_2)]^2 f_{iXj}(a_1, x, a_2)\mathrm{d}(a_1, x, a_2)$$

given that all other terms will depend on $f_{iXj} - g_{iXj}$. As $\mathrm{D}g_{iXj}$ converges faster than h_{iXj}, the leading term is the one associated with h^2_{iXj}. The result follows then from the general tools provided by Aït-Sahalia (1994).

PROOF OF PROPOSITION 2: The conditions imposed by Aït-Sahalia (1994) are such that the functional Taylor expansion under consideration is also valid in the double array case $(d_{i,n}, X_{i,n}, d_{j,n})$. It thus ensues that under $H_1^{[n]}$

$$\hat{\lambda}_n - \frac{b_n^{1/2}}{\hat{\sigma}_\Lambda} \sum_{k=1}^{n-i+j} \left[f_{iXj}(d_{k+i-j,n}, X_{k+i-j,n}, d_{k,n}) - g_{iXj}(d_{k+i-j,n}, X_{k+i-j,n}, d_{k,n})\right]^2$$

converges weakly to a standard normal distribution under $f^{[n]}$. The result then follows by noting that $\hat{\sigma}_\Lambda \xrightarrow{p^{[n]}} \sigma_\Lambda$ and

$$\begin{aligned}
\Lambda_{f^{[n]}} &= E\left[f^{[n]}(d_{i,n}, X_{i,n}, d_{j,n}) - g^{[n]}(d_{i,n}, X_{i,n}, d_{j,n})\right]^2 + O_p\left(n^{-1/2}\right) \\
&= \epsilon_n^2 E\left[\ell^2(d_{i,n}, X_{i,n}, d_{j,n})\right] + o_p\left(n^{-1} b_n^{-1/2}\right) \\
&= n^{-1} b_n^{-1/2} \ell_2 + o_p\left(n^{-1} b_n^{-1/2}\right).
\end{aligned}$$

References

Aït-Sahalia, Yacine, 1994, The delta method for nonparametric kernel functionals, Graduate School of Business, University of Chicago.

Aït-Sahalia, Yacine, 1997, Do interest rates really follow continuous-time Markov diffusions?, Graduate School of Business, University of Chicago.

Amaro de Matos, João, and João S. Rosário, 2000, The equilibrium dynamics for an endogeneous bid-ask spread in competitive financial markets, Universidade Nova de Lisboa.

Baghehot, Walter, pseudonym, 1971, The only game in town, *Financial Analysts Journal* 27, 12–14.

Dawid, A. Philip, 1979, Conditional independence in statistical theory, *Journal of the Royal Statistical Society B* 41, 1–31.

Easley, David, and Maureen O'Hara, 1992, Time and the process of security price adjustment, *Journal of Finance* 47, 577–605.

Fernandes, Marcelo, and Renato G. Flôres, 1999, Non-parametric tests for the Markov property, European University Institute and Getúlio Vargas Foundation.

Fernandes, Marcelo, and Joachim Grammig, 2000, Non-parametric specification tests for conditional duration models, Working Paper ECO 2000/4 European University Institute.

Giot, Pierre, 1999, Time transformations, intraday data and volatility models, CORE Discussion Paper 9944 Université Catholique de Louvain.

Goodhart, Charles A. E., and Maureen O'Hara, 1997, High frequency data in financial markets: Issues and applications, *Journal of Empirical Finance* 4, 73–114.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

O'Hara, Maureen, 1995, *Market Microstructure Theory.* (Blackwell Publishers Cambridge).

Phillips, Peter C. B., and Pierre Perron, 1988, Testing for a unit root in time series regression, *Biometrika* 75, 335346.

Platen, Eckhard, and Rolando Rebolledo, 1996, Principles for modelling financial markets, *Journal of Applied Probability* 31, 601–613.

Silverman, Bernard W., 1986, *Density Estimation for Statistics and Data Analysis.* (Chapman and Hall London).

Snell, Andy, and Ian Tonks, 1998, Testing for asymmetric information and inventory control effects in market maker behaviour on the London Stock Exchange, *Journal of Empirical Finance* 5, 1–25.

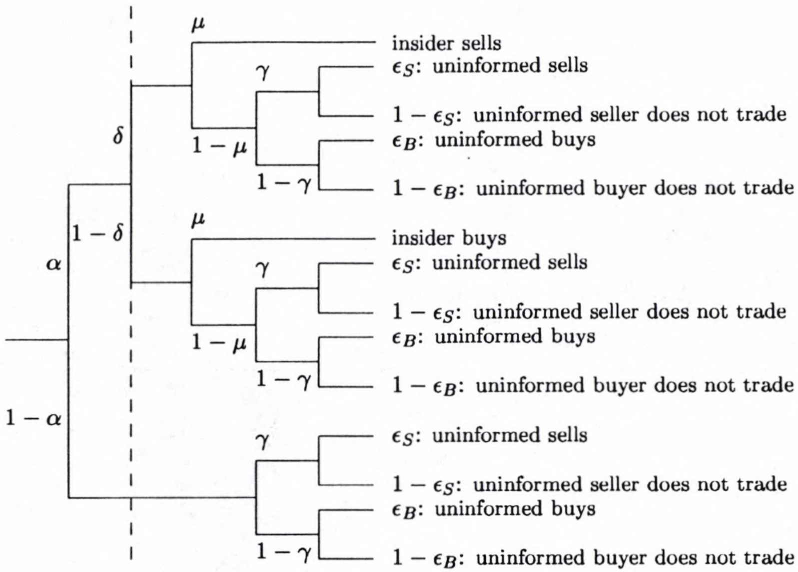

Figure 1 — Tree diagram of the trading process

Notation: α is the probability of an information event, δ is the probability of a low signal, μ is the probability that a trade comes from an informed trader, γ is the probability that an uninformed trader is a seller, $1-\gamma$ is the probability that an uninformed trader is a buyer, ϵ_S is the probability that the uninformed trader will sell, and ϵ_B is the probability that the uninformed trader will buy. Nodes to the left of the dotted line occur only at the beginning of the trading day; nodes to the right occur at each trading interval.

Table I

Phillips and Perron's (1988) unit root tests

stock		sample size	truncation lag	test statistic
Boeing	ask	6,317	10	-1.6402
	bid	6,317	10	-1.6655
	spread	6,317	10	-115.3388
Coca-Cola	ask	3,823	8	-2.1555
	bid	3,823	8	-2.1615
	spread	3,823	8	-110.2846
Disney	ask	5,801	9	-1.2639
	bid	5,801	9	-1.2318
	spread	5,801	9	-112.1909
Exxon	ask	6,009	9	-0.6694
	bid	6,009	9	-0.6405
	spread	6,009	9	-121.8439
IBM	ask	15,124	12	-0.2177
	bid	15,124	12	-0.2124
	spread	15,124	12	-163.0558

Both ask and bid prices are in logs, whereas the spread refers to the difference of the logarithms of the ask and bid prices. The truncation lag ℓ of the Newey and West's (1987) heteroskedasticity and autocorrelation consistent estimate of the spectrum at zero frequency is based on the automatic criterion $\ell = [4(T/100)^{2/9}]$, where $[z]$ denotes the integer part of z.

Table II

Nonparametric tests of the Markov property

stock	duration		adjusted duration	
	$\hat{\lambda}_n$	p-value	$\hat{\lambda}_n$	p-value
Boeing	3.2935	(0.0005)	5.0327	(0.0000)
Coca-Cola	19.4971	(0.0000)	17.9364	(0.0000)
Disney	-2.5783	(0.9950)	-1.6367	(0.9491)
Exxon	-0.5484	(0.7083)	1.3966	(0.0813)
IBM	21.2064	(0.0000)	17.9247	(0.0000)

Adjusted durations refer to the correction for time-of-day effects.

EUI WORKING PAPERS

EUI Working Papers are published and distributed by the European University Institute, Florence

Copies can be obtained free of charge - depending on the availability of stocks - from:

The Publications Officer
European University Instiute
Badia Fiesolana
I-50016 San Domenico di Fiesole (FI)
Italy

Please use order form overleaf

Publications of the European University Institute

To The Publications Officer
European University Institute
Badia Fiesolana
I-50016 San Domenico di Fiesole (FI) - Italy
Fax: +39-055-4685 636
e-mail: publish@datacomm.iue.it
http://www.iue.it

From Name ....

Address ....

....

....

....

....

- ❑ Please send me a complete list of EUI Working Papers
- ❑ Please send me a complete list of EUI book publications
- ❑ Please send me the EUI brochure Academic Year 2000/2001

Please send me the following EUI Working Paper(s):

Dept, n°, author ....

Title: ....

Dept, n°, author ....

Title: ....

Dept, n°, author ....

Title: ....

Dept, n°, author ....

Title: ....

Date ....

Signature ....

Working Papers of the Department of Economics Published since 1999

ECO No. 99/1
Jian-Ming ZHOU
How to Carry Out Land Consolidation - An International Comparison

ECO No. 99/2
Nuala O'DONNELL
Industry Earnings Differentials in Ireland: 1987-1994

ECO No. 99/3
Ray BARRELL/Rebecca RILEY
Equilibrium Unemployment and Labour Market Flows in the UK

ECO No. 99/4
Klaus ADAM
Learning while Searching for the Best Alternative

ECO No. 99/5
Guido ASCARI/Juan Angel GARCIA
Relative Wage Concern and the Keynesian Contract Multiplier

ECO No. 99/6
Guido ASCARI/Juan Angel GARCIA
Price/Wage Staggering and Persistence

ECO No. 99/7
Elena GENNARI
Estimating Money Demand in Italy: 1970-1994

ECO No. 99/8
Marcello D'AMATO/Barbara PISTORESI
Interest Rate Spreads Between Italy and Germany: 1995-1997

ECO No. 99/9
Søren JOHANSEN
A Small Sample Correction for Tests of Hypotheses on the Cointegrating Vectors

ECO No. 99/10
Søren JOHANSEN
A Bartlett Correction Factor for Tests on the Cointegrating Relations

ECO No. 99/11
Monika MERZ/Axel SCHIMMELPFENNIG
Career Choices of German High School Graduates: Evidence from the German Socio-Economic Panel

ECO No. 99/12
Fragiskos ARCHONTAKIS
Jordan Matrices on the Equivalence of the $I(1)$ Conditions for VAR Systems

ECO No. 99/13
Étienne BILLETTE de VILLEMEUR
Sequential Decision Processes Make Behavioural Types Endogenous

ECO No. 99/14
Günther REHME
Public Policies and Education, Economic Growth and Income Distribution

ECO No. 99/15
Pierpaolo BATTIGALLI/ Marciano SINISCALCHI
Interactive Beliefs and Forward Induction

ECO No. 99/16
Marco FUGAZZA
Search Subsidies vs Hiring Subsidies: A General Equilibrium Analysis of Employment Vouchers

ECO No. 99/17
Pierpaolo BATTIGALLI
Rationalizability in Incomplete Information Games

ECO No. 99/18
Ramon MARIMON/Juan Pablo NICOLINI/Pedro TELES
Competition and Reputation

ECO No. 99/19
Ramon MARIMON/Fabrizio ZILIBOTTI
Employment and Distributional Effects of Restricting Working Time

*out of print

ECO No. 99/20
Leonor COUTINHO
Euro Exchange Rates: What Can Be Expected in Terms of Volatility?

ECO No. 99/21
Bernard FINGLETON
Economic Geography with Spatial Econometrics: A 'Third Way' to Analyse Economic Development and 'Equilibrium', with Application to the EU Regions

ECO 99/22
Mike ARTIS/
Massimiliano MARCELLINO
Fiscal Forecasting: The Track Record of the IMF, OECD and EC

ECO 99/23
Massimo MOTTA/Michele POLO
Leniency Programs and Cartel Prosecution

ECO 99/24
Mike ARTIS/Hans-Martin KROLZIG/
Juan TORO
The European Business Cycle

ECO 99/25
Mathias HOFFMANN
Current Accounts and the Persistence of Global and Country-Specific Shocks: Is Investment Really too Volatile?

ECO 99/26
Mathias HOFFMANN
National Stochastic Trends and International Macroeconomic Fluctuations: The Role of the Current Account

ECO 99/27
Gianmarco I.P. OTTAVIANO/
Jacques-François THISSE
Integration, Agglomeration and the Political Economics of Factor Mobility

ECO 99/28
Gianmarco I.P. OTTAVIANO
Ad usum delphini: A Primer in 'New Economic Geography'

ECO 99/29
Giorgio BASEVI/
Gianmarco I.P. OTTAVIANO
The District Goes Global: Export vs FDI

ECO 99/30
Hans-Martin KROLZIG/Juan TORO
A New Approach to the Analysis of Shocks and the Cycle in a Model of Output and Employment

ECO 99/31
Gianmarco I.P. OTTAVIANO/
Jacques-François THISSE
Monopolistic Competition, Multiproduct Firms and Optimum Product Diversity

ECO 99/32
Stefano MANZOCCHI/
Gianmarco I.P. OTTAVIANO
Outsiders in Economic Integration: The Case of a Transition Economy

ECO 99/33
Roger E.A. FARMER
Two New Keynesian Theories of Sticky Prices

ECO 99/34
Rosalind L. BENNETT/
Roger E.A. FARMER
Indeterminacy with Non-Separable Utility

ECO 99/35
Jess BENHABIB/Roger E.A. FARMER
The Monetary Transmission Mechanism

ECO 99/36
Günther REHME
Distributive Policies and Economic Growth: An Optimal Taxation Approach

ECO 99/37 - Ray BARRELL/
Karen DURY/Ian HURST
Analysing Monetary and Fiscal Policy Regimes Using Deterministic and Stochastic Simulations

ECO 99/38
Chiara FUMAGALLI/Massimo MOTTA
Upstream Mergers, Downstream Mergers, and Secret Vertical Contracts

ECO 99/39
Marcel FRATZSCHER
What Causes Currency Crises: Sunspots, Contagion or Fundamentals?

ECO 99/40
Simone BORGHESI
Intergenerational Altruism and Sustainable Development

*out of print

ECO 99/41
Jeni KLUGMAN/Alexandre KOLEV
The Role of the Safety Net and the Labour Market on Falling Cash Consumption in Russia: 1994-96. A Quintile-Based Decomposition Analysis

ECO 99/42
Günther REHME
Education, Economic Growth and Personal Income Inequality Across Countries

ECO 99/43
Günther REHME
Why are the Data at Odds with Theory? Growth and (Re-)Distributive Policies in Integrated Economies

ECO 99/44 - Norbert WUTHE
Exchange Rate Volatility: The Impact of Learning Behaviour and the Institutional Framework - A Market Microstructure Approach

* * *

ECO 2000/1
Marcelo FERNANDES
Central Limit Theorem for Asymmetric Kernel Functionals

ECO 2000/2
Ray BARRELL/Álvaro M. PINA
How Important are Automatic Stabilizers in Europe? A Stochastic Simulation Assessment

ECO 2000/3
Karen DURY/ Álvaro M. PINA
Fiscal Policy in EMU: Simulating the Operation of the Stability Pact

ECO 2000/4
Marcelo FERNANDES/Joachim GRAMMIG
Non-parametric Specification Tests for Conditional Duration Models

ECO 2000/5
Yadira GONZÁLEZ DE LARA
Risk-sharing, Enforceability, Information and Capital Structure

ECO 2000/6
Karl H. SCHLAG
The Impact of Selling Information on Competition

ECO 2000/7
Anindya BANERJEE/Bill RUSSELL
The Relationship Between the Markup and Inflation in the G7 Plus One Economies

ECO 2000/8
Steffen HOERNIG
Existence and Comparative Statics in Heterogeneous Cournot Oligopolies

ECO 2000/9
Steffen HOERNIG
Asymmetry and Leapfrogging in a Step-by-Step R&D-race

ECO 2000/10
Andreas BEYER/Jurgen A. DOORNIK/David F. HENDRY
Constructing Historical Euro-Zone Data

ECO 2000/11
Norbert WUTHE
Exchange Rate Volatility's Dependence on Different Degrees of Competition under Different Learning Rules - A Market Microstructure Approach

ECO 2000/12
Michael EHRMANN
Firm Size and Monetary Policy Transmission - Evidence from German Business Survey Data

ECO 2000/13
Günther REHME
Economic Growth and (Re-)Distributive Policies: A Comparative Dynamic Analysis

ECO 2000/14
Jarkko TURUNEN
Leaving State Jobs in Russia

ECO 2000/15
Søren JOHANSEN
A Small Sample Correction of the Test for Cointegrating Rank in the Vector Autoregressive Model

ECO 2000/16
Martin ZAGLER
Economic Growth, Structural Change, and Search Unemployment

*out of print

ECO 2000/17
Martin ZAGLER
Aggregate Demand, Economic Growth, and Unemployment

ECO 2000/18
Martin ZAGLER
Growth Externalities, Unions, and Long-term Wage Accords

ECO 2000/19
João AMARO DE MATOS/
Marcelo FERNANDES
Market Microstructure Models and the Markov Property

*out of print